

Chapter 6

***In Vitro* Selection of Peptides Targeting Human Telomerase RNA**

Abstract

Telomerase is an attractive anti-cancer target since it is expressed in most tumor cells but not in most normal cells. It is a ribonucleoprotein, composed of telomerase RNA and telomerase reverse transcriptase. We have targeted two functionally vital domains of telomerase RNA, P6.1 and CR7, using mRNA display. From peptide libraries of 10 or 100 trillion members, we have isolated L- and D-peptides that bind with low nanomolar affinity and are highly specific for their cognate RNA target. These peptides are highly basic and are predicted to form α -helices, however non-electrostatic interactions contribute to overall peptide binding. Our data argue that very few random peptides bind RNA with high affinity and specificity. We expect that the peptides we have selected should bind in the context of telomerase RNA and will inhibit telomerase activity.

Introduction

Telomerase is a reverse transcriptase that is responsible for maintaining the telomeres at the ends of eukaryotic chromosomes (1). It is minimally composed of telomerase RNA (TR) and a protein component, telomerase reverse transcriptase (TERT), but other accessory proteins are known to associate with the telomerase complex (2). The enzyme utilizes an RNA template provided by TR to add d(TTAGGG) repeats to the ends of telomeres, preventing the progressive telomere shortening associated with DNA replication. Telomere shortening leads to cellular senescence and growth arrest, but cell crisis and death result if cells continue to divide after senescence should have been reached.

Roughly 85-90% of all human cancer cells constitutively express telomerase (3). Expression of telomerase is one of three events required for the oncogenic transformation of normal human cells (4). Most normal cells show little expression of telomerase, with the exception of stem and germ-line cells. Telomere lengths of normal cells are generally longer than those of tumor cells (5). Telomerase, therefore, represents an attractive target for anti-tumor specific therapy.

Several strategies currently exist for inhibiting telomerase. One class of compounds inhibits telomerase by stabilizing G-quartet structures, preventing telomerase from acting on the G-rich telomeres (6, 7). Other groups have used reverse transcriptase inhibitors to target hTERT (8). Lastly, several compounds that basepair with human TR (hTR) have been used to target the template region (9-11) and prevent assembly of the human TERT (hTERT)/hTR complex (12).

No group has targeted specific structural elements of hTR, possibly because of the difficulty in designing ligands to structural, rather than sequence elements. Targeting structural elements has the potential to be more specific than targeting sequence-specific elements. It is also more difficult for drug resistance to develop against structural-based compounds since there are few mutations that would prevent ligand binding but retain interactions with natural binding partners.

We have previously used mRNA display to isolate novel, high affinity, RNA binding peptides (13-15). mRNA display enables the selection of peptides by covalently coupling the peptides to their encoding mRNA (Figure 6.1), allowing the peptide sequence to be recovered by amplification of the RNA (16). Libraries containing up to 100 trillion independent sequences (10^{14}) can be constructed by mRNA display, the

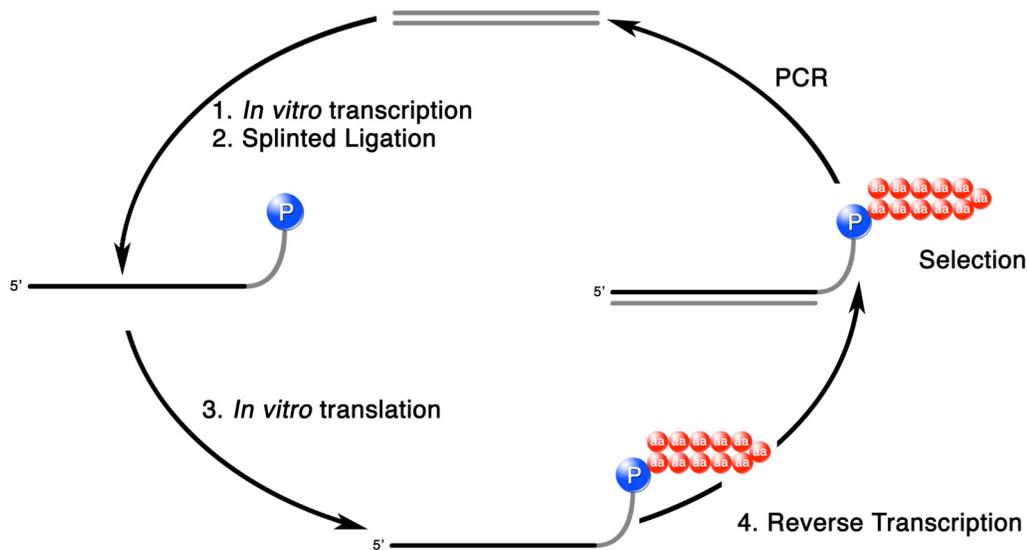


Figure 6.1. An mRNA display selection cycle. A double stranded DNA library is transcribed using T7 RNA polymerase [1], and the resulting mRNA ligated to a synthetic oligonucleotide containing puromycin [2]. *In vitro* translation of the ligated product results in attachment of a peptide to its encoding mRNA [3]. Reverse transcription generates a cDNA/mRNA hybrid [4], which is used in affinity selection. PCR generates an enriched pool which is used in further cycles of selection.

largest of any peptide library. We sought to extend our original methodology and develop methods for the selection of RNA binding peptides targeting RNAs for which no natural ligand is known. We chose to target telomerase RNA since hTR-binding peptides would be useful tools to study telomerase as well as potential drug candidates.

Our overall design strategy to isolate novel RNA binding ligands is outlined in (Figure 6.2). Based on the RNA secondary structure of telomerase, we identified regions important for activity by their high sequence conservation or by loss of function when mutagenized, and constructed RNA targets based on these domains. The RNA targets were structurally

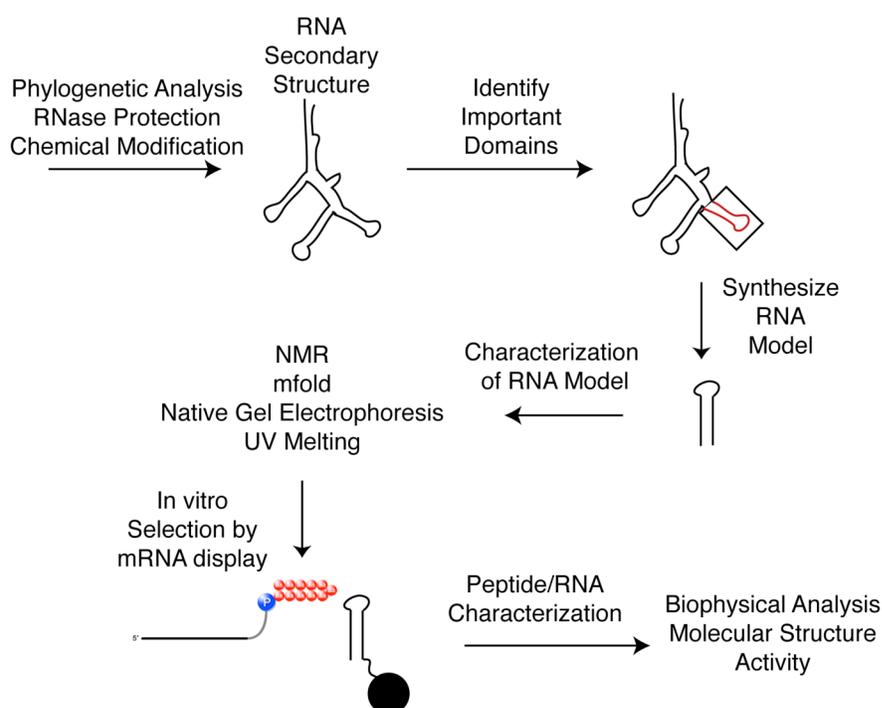


Figure 6.2. Strategy for selection of novel RNA binding peptides. Using a phylogenetic or experimentally derived RNA secondary structure, domains important for RNA activity are identified (shown in red). Small RNA models based on important domains are synthesized and characterized using NMR and native gel electrophoresis to insure that the RNA is folded and monomeric. A biotinylated target is synthesized and used as a target in an mRNA display selection. Peptide sequences from the selection are synthesized, tested for activity, as well as characterized biophysically and structurally.

characterized using NMR and native gel electrophoresis to confirm the RNA is folded and monomeric. We then synthesized biotinylated RNAs (both natural and unnatural, see below) and used them as targets in separate mRNA display selections. Peptides resulting from the selection were characterized using a variety of biophysical and structural techniques and will be tested for activity both *in vivo* and *in vitro*. This selection strategy should be applicable to isolate peptides that bind any RNA target.

Based on our previous experience, we expect that at least some of the peptides we select will have an affect on telomerase activity (15). RNA binding peptides often contain arginine rich domains (17) and such highly basic sequences can cross cellular membranes (18, 19). However, peptides are usually degraded by biological systems, reducing their overall effectiveness. We aim to select peptides that have better stability in biological systems using a reflection selection strategy (20, 21) (Figure 6.3). Selection

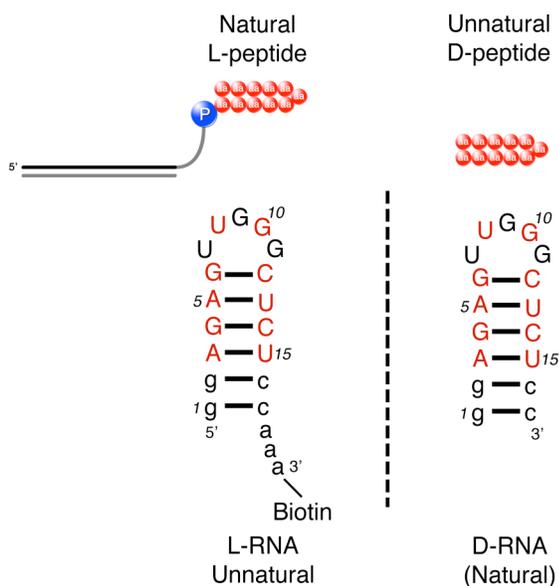


Figure 6.3. Reflection selection strategy. An unnatural (opposite chirality) L-RNA is synthesized and used as a target in an mRNA display selection using natural L-amino acids. Winning sequences are synthesized as D-peptides, which bind to natural D-RNA and are resistant to protease degradation.

is performed against an unnatural RNA target composed of L-nucleotides instead of D-nucleotides resulting in L-peptides that bind L-RNA. D-peptides of the same sequences should bind the natural D-RNA, resulting in peptides that are resistant to degradation by proteases (22).

Results and Discussion

RNA Target Determination

The ideal RNA target would have the following features: (i) Be important for telomerase function so that peptides that bind it would inhibit telomerase (ii) Be small enough to be synthesized by chemical methods so that unnatural bases could be introduced (e.g., 2-aminopurine, a fluorescent base used for biophysical studies) and since smaller RNAs are more amenable to structure determination by NMR (iii) Be stable and well folded (iv) Be monomeric.

Although we could theoretically target the full-length telomerase RNA, we were concerned that hTR would not be an ideal target for a number of reasons. First, full-length hTR may not fold correctly in the absence of hTERT, resulting in peptides that would bind non-physiologically relevant structures (23). A large RNA also has more potential for interaction with the nucleic acid component of mRNA-protein fusions resulting in selection for RNA-RNA or RNA-DNA interactions, instead of peptide-RNA interactions. Lastly, once peptides were isolated after selection, determining where individual peptides bind would be complicated since there are presumably many sites on hTR that could serve as targets for our peptides.

The phylogenetically determined secondary structure of vertebrate telomerase RNA (Figure 6.) contains eight conserved regions (CR1-8) (24, 25). CR1 corresponds to the template region and has been targeted by a number of antisense compounds (26). The pseudoknot CR2/CR3 domain and the CR4/CR5 domain are both important for

catalytic activity. The CR6/CR8 (boxH/ACA) and CR7 domains have been implicated in RNA processing and localization (27).

We focused our initial efforts on the CR4/CR5 domain and the CR7 domain, which are important for telomerase function. Construction of model RNA targets based on these domains was straightforward since both domains contain RNA hairpins, the targets of other selections using mRNA display (13, 14). The CR1 domain was also a potential target, however we feared it would be difficult for small peptides to recognize this single-stranded domain.

The CR4/CR5 domain encompasses almost 60 nt, with many highly conserved nucleotides located in the CR5 region. Part of the CR5 region folds into a small (13 nt), universally conserved hairpin, designated P6.1 (25). P6.1 is critical for telomerase activity; disrupting the stem or mutating the highly conserved U307 or G309 abolishes function in mouse telomerase. However, it is likely that these nucleotides would also be important in human telomerase because of their strict conservation. The P6.1 loop adopts a backbone conformation similar to UNCG tetraloops (28), with one base (G309) extruded (Figure 6.). In the P6.1 NMR structure (29), U306 and G310 form a GU wobble pair and the bases in the loop are highly exposed to solvent, suggesting that they interact with hTERT or other regions of hTR (29, 30).

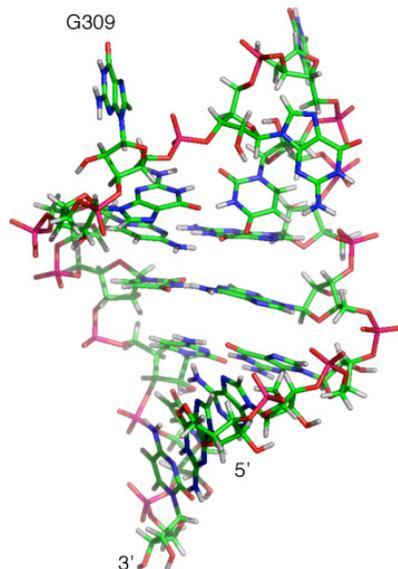


Figure 6.5. NMR structure of the P6.1 hairpin. The P6.1 pentaloop adopts a UNCG-like tetraloop fold, extruding G309 from the structure (29). The 5' and 3' ends of the hairpin are indicated.

The CR7 domain is a 30 nt domain that forms a hairpin containing an internal bulge. Many highly conserved bases are located in the CR7 terminal hairpin loop (U407-A422) and mutations or deletions in this region prevent hTR accumulation *in vivo* (31). The CR7 loop contains a self-complementary sequence (5'-AGCU-3') that could potentially form kissing complexes (32). The ability of hTR to dimerize may play a role *in vivo* as functional telomerase complexes are dimeric (33).

We designed and synthesized two RNA hairpins based on the CR4/CR5 P6.1 hairpin (A302-C314) and the terminal CR7 hairpin (U407-A422)(Figure 6.). One and two additional GC pairs were added to the CR7 and P6.1 stems, respectively, to stabilize the hairpins and allow the RNA to be synthesized by *in vitro* transcription (34).

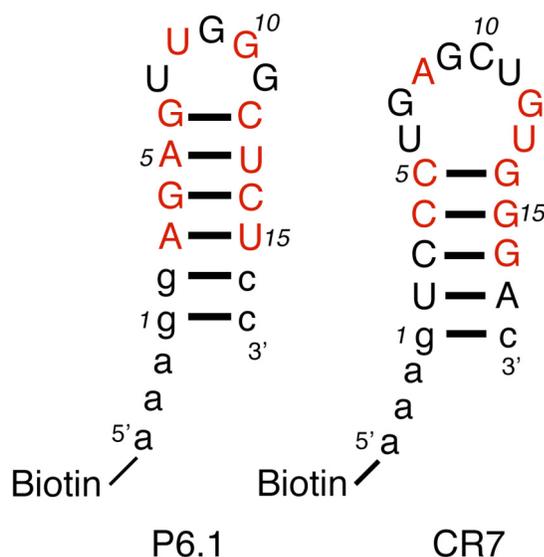


Figure 6.6. Sequence and secondary structure of telomerase RNA models. Highly conserved positions are shaded red. RNA models used for native gel electrophoresis and NMR lack the 5' biotin and rArArA spacer. Nucleotides added for stability and as spacers are depicted in lower case.

Characterization of Designed Sequences

Native gel electrophoresis of the RNA hairpins showed significant dimerization of the RNAs (Figure 6.). However, annealing by rapid heating and cooling of the sample showed that more than 90% of the RNA could be shifted to the monomeric form (Figure 6.). P6.1 remained monomeric for more than 4 hours while CR7 quickly dimerized, although the dimer form generally composed less than 10% of the sample (data not shown).

Imino protons in 1D ^1H -NMR spectra provide useful information about RNA secondary structure. Generally, a base-paired imino proton is protected from exchange with solvent resulting in a peak in the NMR spectrum. NMR of the annealed RNA hairpins showed reasonably sharp, well defined peaks in the imino region implying that

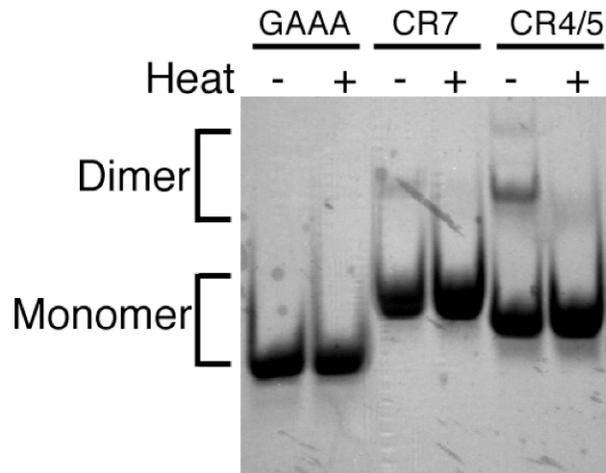


Figure 6.7. Native gel analysis of telomerase RNA targets. RNAs were annealed (+ heat lanes) by heating to 90 °C and rapidly cooling on ice. A control hairpin derived from *boxB* RNA containing the sequence 5'-GAAA-3' in the loop is shown as a monomeric control. Monomeric and dimeric species are indicated.

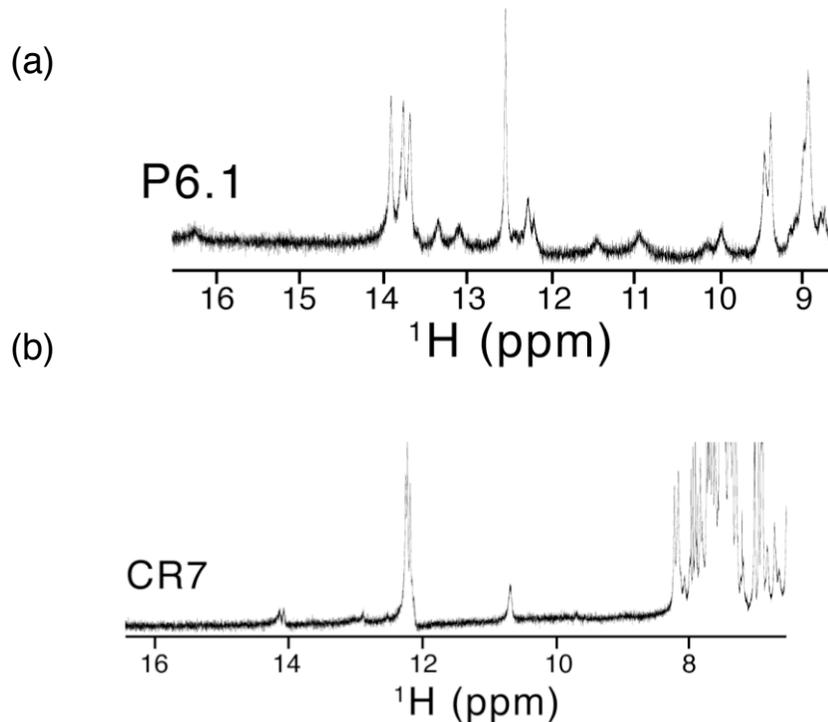


Figure 6.8. NMR spectra of RNA targets. (a) D-P6.1 hairpin was heated and cooled rapidly before the NMR experiment. The imino region of the spectrum is shown. Peaks near 14 p.p.m. are most likely from G nucleotides while the peak at 12.5 p.p.m. results from a U nucleotide. Small peaks near 11 and 13 p.p.m. are possibly from RNA dimerization. (b) NMR of D-CR7. The sample was prepared as in (a).

the hairpins are folded (Figure 6.8). However, several small peaks in the imino region were observed, suggesting the presence of a small amount of dimeric RNA. The high concentrations of hairpin required for NMR probably enhanced dimerization even though the hairpins were annealed before the spectra were taken.

These data suggested that the P6.1 and CR7 models were both folded and mostly monomeric assuring us that these RNA would be good targets for selection. Moreover, the low concentrations of target required for selection ($\sim 0.2 \mu\text{M}$) would further reduce the dimer population in our samples. Thus, we synthesized two biotinylated D-RNA targets corresponding to D-P6.1 and D-CR7. We also synthesized an L-P6.1 target for the reflection selection strategy.

Library Design

There are no known natural ligands for the P6.1 hairpin on which to base the design of our library (13). We therefore constructed two random libraries that differed on the degree of randomization. The first library contained 27 random amino acids (X27) and has been previously used to discover peptides that bind Methuselah, a *Drosophila melanogaster* G-protein coupled receptor (B. Ja. and R.W. Roberts, in preparation). We reasoned that the totally random library would sample sequence space evenly, preventing any bias in the solutions obtained. The library contained a 3' SpeI cloning site, coding for the amino acid sequence TS and a GGLRASAI C-terminal constant region.

The second library was based on the observation that many RNA binding peptides contain arginine-rich domains (arginine-rich motif, ARM) (35). Alignment of a number of ARM sequences (Figure 6.3) showed many sequences possessed an RRXRR sequence.

We designed a random library containing a central RRXRR motif where each Arg was doped at ~50% (the probability of Arg at all four positions is 12.5%). The library also incorporated a 3' NgoMIV cloning site, coding for the amino acid sequence AG, a C-terminal tyrosine to allow for easy quantitation of selected peptides, and a QLRNSCA constant region (13).

λ N ₍₁₋₂₂₎	MDAQTRRRERRAEKQAQWKAAN
P22 N ₍₁₄₋₃₀₎	NAKTRHERRRKLAIER
ϕ 21 N ₍₁₂₋₃₉₎	TAKTRYKARRAELIAERR
HIV-Rev ₍₃₄₋₅₀₎	RQARRNRRRWRRERQR
HIV-TAT ₍₄₉₋₅₇₎	RKKRRQRRR
BMV Gag ₍₇₋₂₅₎	KMTRAQRRAAARRNRWTAR
CCMV Gag ₍₇₋₂₅₎	KLTRAQRRAAARKNKRNTR
Yeast PRP6 ₍₁₂₉₋₁₄₄₎	TRRNKRNRRIQEQLNRK
Human U2AF ₍₁₄₂₋₁₅₃₎	SQMTRRARRLYV
L16 ₍₅₁₋₆₃₎	RRAMSRKFRNSK
S7 ₍₉₁₋₉₉₎	KTKLERRNK
HTLV-II Rex ₍₄₋₁₆₎	TRRQRTRRARRNR
FHV coat ₍₃₅₋₄₉₎	RRRRNRTRRNRRRVR
Consensus	RRXRR

Figure 6.3. Alignment of several arginine rich peptides. Adapted from (35). Arginine rich regions are often of the sequence RRXRR.

D-CR7 Selection

Utilizing the RRXRR library, we synthesized a pool of mRNA protein fusions containing 100 trillion (10^{14}) independent sequences and sieved the pool against immobilized D-CR7. We performed a total of fifteen rounds of selection, alternating matrices every few rounds to avoid selecting matrix-binding sequences (Figure 6.4). The round 15 pool bound reasonably well, but also contained modest affinity to the matrix

with no target (data not shown). Attempts to increase binding affinity by increasing binding stringency or decrease nonspecific binding by using different immobilization matrices and by preclearing the library were unsuccessful. The round 15 library was sequenced (Figure 6.5) yielding one sequence that dominated the pool. The winning sequence contains a two-base deletion in the 3' NgoMIV restriction site causing a shift into the +3 reading frame.

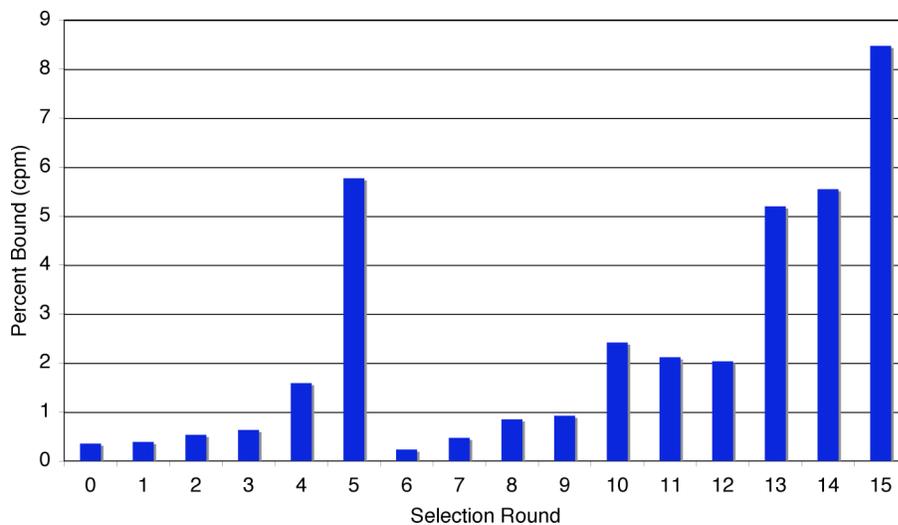


Figure 6.4. Binding of the round 0-15 pools to immobilized D-CR7. Binding percentages are corrected for matrix binding.

Clone	Peptide Sequence			
	1	10	20	30
Pool	•	•	•	•
	MXXXXXrrrXrrXXXXXXXXXXAGYQLRNSCA			
CR7_15J1	MFLIYFDRVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J2	MFLIYFDHVRRRMKIDFIPSGPLPAQKQLR			
CR7_15J3	MFLIYFDRVRRRMKIDFIPSGPLPAQKQLR			
CR7_15J4	MFVIYFDRVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J5	MFLIYFDHVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J6	MFLIYFDRVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J7	MFLIYFDHVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J8	MFVIYFDRVRRRMKIDFIPSGSLPAQKQLR			
CR7_15J9	MFVIYFDRVRRRMKIDFIP T GSLPAQKQLR			
CR7_15J10	MFVIYFDRVR Q RMKIDFIPSGSLPAQKQLR			

Figure 6.5. Peptide sequences from round 15J. Mutations are highlighted in bold text. The RRXRR pool sequence is shown as a reference.

A nonmutated winning sequence (15J3) and two sequences representing the common mutations R8H (15J2) and L3V (15J8) were selected for further study. Each clone was amplified from plasmid DNA, synthesized as a ³⁵S-labeled fusion, and used in an *in vitro* binding assay against immobilized CR7. The results (Figure 6.6) show that 15J2, 15J3, and 15J8 bind CR7 with similar affinity, suggesting that the R8H and L3V mutations do not affect the binding affinity of the peptides. However, the 15J3 sequence shows some affinity toward the matrix, explaining why attempts to reduce nonspecific binding failed. The 15J3 sequence also shows low net binding to two other immobilized RNAs (D-P6.1 and D-HCVIIIId RNA).

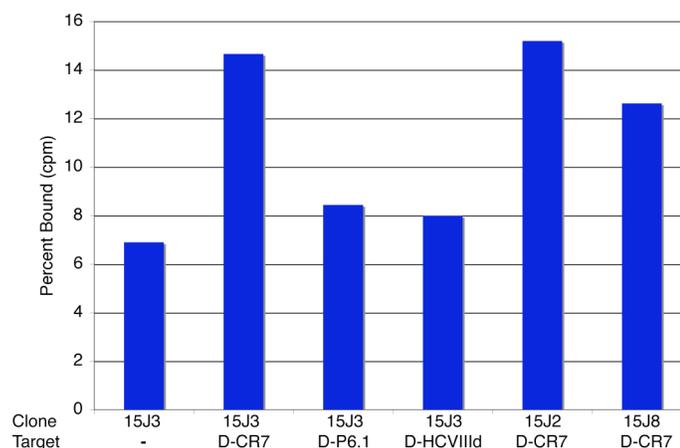


Figure 6.6. Binding of different round 15J clones to immobilized RNA targets. Individual clones corresponding to 15J2, 15J3, and 15J8 were synthesized as ^{35}S -labeled fusions and tested for binding to immobilized D-CR7, DHCVIIIId, and D-P6.1. A no target control is shown for the 15J3 sequence (-).

DP6.1 Selection

We performed a similar selection targeting immobilized D-P6.1 using the RRXRR library. Using a 100 trillion-member library (10^{14} sequences), we performed a total of eight rounds of selection against immobilized D-P6.1 (Figure 6.7). The matrix was switched to streptavidin magnetic beads at round five in order to reduce matrix-binding sequences. Increasing binding stringency at round eight resulted in no further increase in binding, and the round eight pool was sequenced (Figure 6.8). Once again, a single sequence (termed 8H1) dominated the pool. 8H1 and a truncated variant changing the C-terminal constant region (QLRNCSA) to GGGG were amplified from plasmid DNA, synthesized as ^{35}S -labeled fusions, and tested for binding against immobilized D-P6.1. Figure 6.9 shows that both 8H1 and truncated 8H1 bind to D-P6.1, however truncated 8H1 possesses higher affinity than full-length 8H1 demonstrating that the QLRNCSA constant region is not required for, and may even hinder binding.

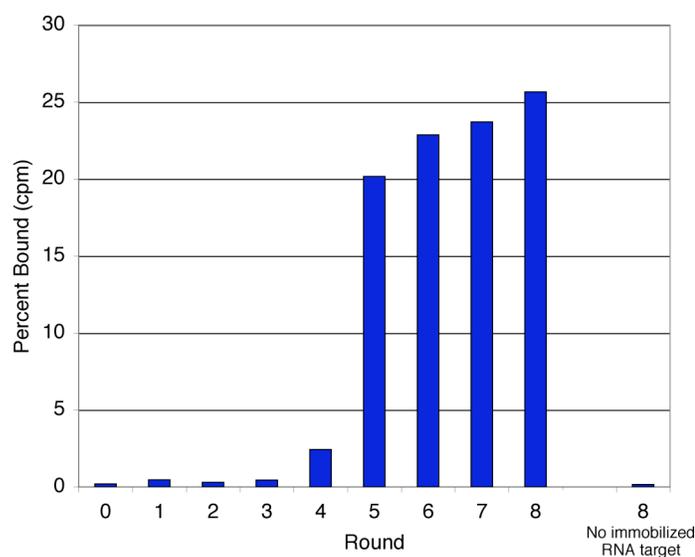


Figure 6.7. Binding of rounds 0-8 to immobilized D-P6.1. A no target control from round 8 is shown.

Clone	Peptide Sequence			
	1	10	20	30
Pool	•	•	•	•
	MXXXXXrrrXrrXXXXXXXXXXAGYQLRNCSA			
RRXRR8H.1	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			
RRXRR8H.2	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			
RRXRR8H.3	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			
RRXRR8H.4	TNDARRNRKYLrvKRLRIQKT--YQLRNCSA			
RRXRR8H.6	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			
RRXRR8H.7	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			
RRXRR8H.8	MNDARRNRKYLrvKRLRIQKM--YQLRNCSA			

Figure 6.8. Sequences from round 8H targeting D-P6.1. The starting library is shown for reference. Lowercase (r) positions indicate a 50% probability of Arg. Dashes (-) indicate a deletion. The red period indicates a stop codon. Peptide mutations are highlighted in bold text.

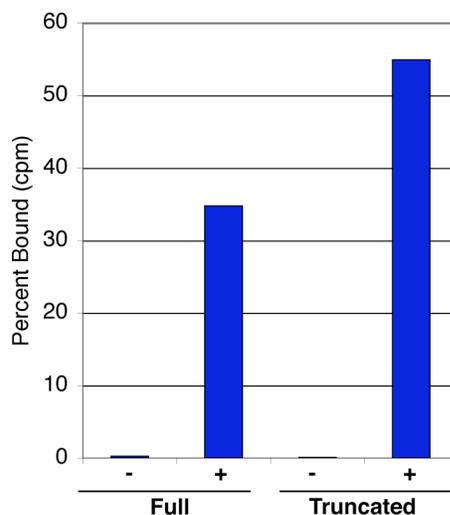


Figure 6.9. Binding of full-length and truncated 8H1. Truncated 8H1 replaces the QLRNSCA constant region with GGGG. Binding reactions containing no target (-) are shown as controls.

We chemically synthesized a peptide based on truncated 8H1 for further biophysical characterization. Attempts to determine a binding constant by a gel shift assay or by titration against 2-aminopurine labeled P6.1 were unsuccessful (data not shown). Preliminary experiments using surface plasmon resonance yield a k_{on} of $2.2 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$, a k_{off} of 0.048 s^{-1} , and a K_{d} of 2.2 nM (Figure 6.10). However, binding under these experimental conditions is mass-transport limited and exacerbated by the fact that only ~25% of immobilized D-P6.1 is functional, indicating that P6.1 may be conformationally heterogeneous.

8H1 is highly homologous to the winning sequence (18B) selected from a 10^{13} -member library targeting D-P6.1 using the X27S library (C. Ueda, unpublished observations). Alignment of 8H1 and 18B (Figure 6.11) shows that ten amino acids are conserved, suggesting these positions are involved in binding. Both sequences are

predicted to form α -helices (36) and the conserved residues are spaced with an $i, i+3$ or $i, i+4$ pattern suggesting they are located on the same face of an α -helix.

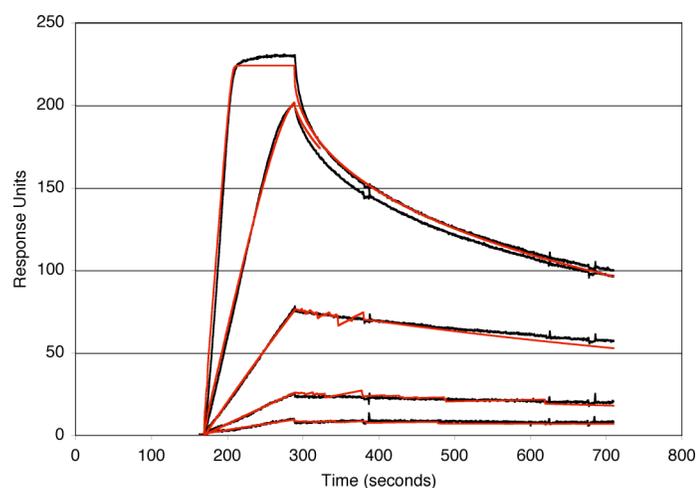


Figure 6.10. The binding of 8H1 studied by surface plasmon resonance. 8H1 was injected at 1.22, 3.67, 11, 33, and 99 nM and bound to D-P6.1 immobilized on a streptavidin sensor chip. Sensorgrams (black) are corrected for drift using a flow cell containing no RNA target. Kinetic fits are shown in red.

8H1	MNDARRNRKYLRVKRLRIQKMYQLRNSCA
18B	MTTSARSLRKYYRVLLRAFKTRPARHVTS <u>SGGLRASAI</u>
Consensus	AR RKY RV+ LR K

Figure 6.11. Alignment of 18B and 8H1 peptides. 18B was selected from the X27S library while 8H1 was selected from the RRXRR library. Constant regions are underlined. Positions common to both peptides are highlighted in bold text. (+) designates the position requires a positive charge (Lys or Arg).

The 8H1 sequence also contains a six-base deletion that destroys the NgoM IV restriction site. The deletion of the NgoM IV restriction site has also been observed in the D-CR7 selection (above) and other selections using the RRXRR library (data not shown). The deletion of the site is unlikely to be due to poor construction of the library as the majority of the sequences in the starting pool contain the NgoM IV site (Figure

6.S1). It is possible that the amino acid sequence (AG) is a site for protease cleavage, however fusion formation of the RRXRR library is comparable to other libraries used for selection. A likely explanation is that the nucleotide sequence of the restriction site (5'-GCCGGC-3') resulted in poor amplification of the template sequence. The NgoM IV restriction site has been used previously (13), however in that experiment, it was not subjected to multiple cycles of selection, as in these experiments. Regardless of the reason for bias, the restriction site should be avoided for future mRNA display libraries.

L-P6.1 Selection

It was unclear structurally how and if an L-peptide would bind to an L-RNA. We thus chose the totally random X27S library to target L-P6.1 since it would not bias the

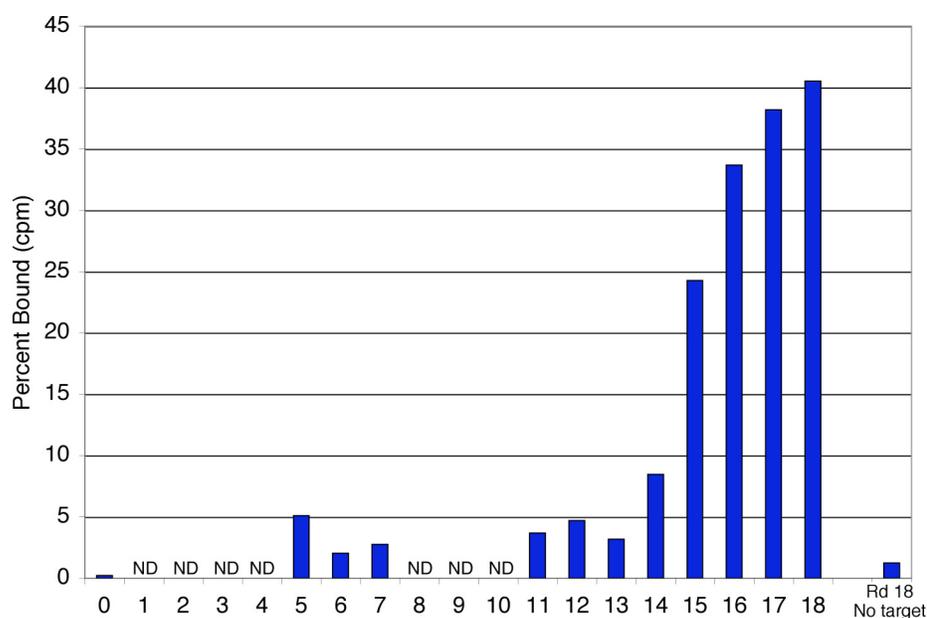


Figure 6.12. Binding of rounds 0-18 to immobilized L-P6.1. (ND) indicates that no binding assay was performed for that particular round. Binding of the round 18 pool to matrix containing no L-P6.1 is shown as a control.

selection. We constructed an mRNA display library totaling ten trillion (10^{13}) independent sequences and sieved the library against immobilized L-P6.1. After 16 rounds of selection, the library bound L-P6.1 very well (Figure 6.12) and we tested the ability of the pool to bind other immobilized RNA targets. Incubation of ^{35}S -labeled fusions with L- versions of *boxB* and CR7 as well as D-P6.1 resulted in little binding to these noncognate targets (Figure 6.13). However, the pool does show binding to an L-HCV IIIId target (data not shown, discussed in Chapter 7).

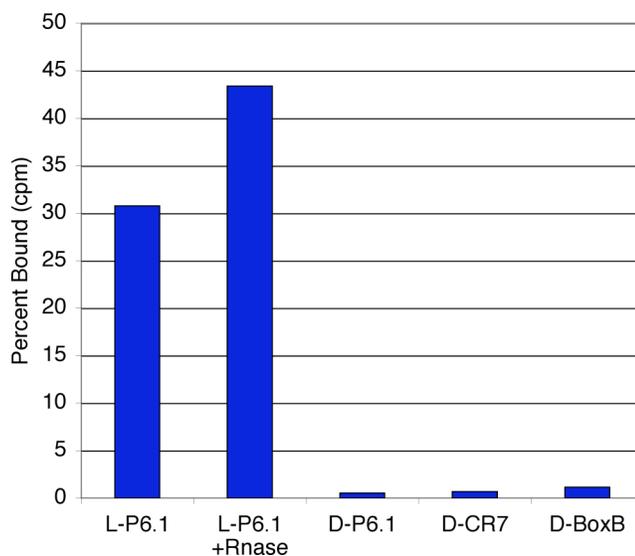


Figure 6.13. The specificity of the round 16 pool of the L-P6.1 selection. ^{35}S -labeled fusions were incubated with different immobilized RNA targets and binding percentages were determined by scintillation counting. A sample was treated with RNase (+RNase) to show that binding is not due to the mRNA component of the fusion.

It was also necessary to test if the mRNA contributed any affinity to the binding interaction. Addition of RNase to the binding reaction shows that the fusions retain binding to the L-P6.1 target ruling out any binding contribution of the mRNA (Figure 6.20). In fact, the RNase-treated fusions show higher binding affinity than the mRNA-

protein fusions, as observed previously (13). These data also demonstrate that the L-enantiomer of RNA is resistant to degradation by RNase.

After two additional rounds of selection, little increase in binding was seen and the pool was sequenced. Figure 6.14 shows that the selected sequences belong to three sequence families with no consensus to each other. One sequence (18C1) composed roughly 60% of all clones (10/17) while another sequence (18C5) composed roughly a third of all sequences (6/17). The third sequence (18C17) was represented only once.

Clone	Peptide Sequence			
	1	10	20	30
Pool	MXXXXXXXXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXXXXXXXXX	XXXXXXXXXXXXXXXXXXXXTSGGLRASAI
18C1, 2, 9, 12, 19	MMDWKRAKLNRLSVRKL	RKYADYF	----	TSGGLRASAI
18C7	MMDWKRAKLSR	SVRKL	RKYADYF	---- A S
18C13	-MDWKRAKLSR	SVRKL	RKYADYF	----TSGGLRASAI
18C11	MTDYKRAKLNLLSVRKL	RKYADYF	----	TSGGLRASAI
18C14	MTACKRAKLNRLSVRKL	RKYADYF	----	TSGGLRASAI
18C18	MMDWRRAKLNRLS	A RKLRKYADYF	----	TRGGLRASAI
18C5, 10	MGYLTPKGRALKRMLDRNRRRKAKSGVT	S		
18C20	MGYLTPKGRALKRMLDRNRRRKVKSGVT	S		
18C4	MGYLTPKGRALKRMLSRNRRREAKSGVT	S		
18C6	MGYLTPKGRALKRMLDRIRRRREAKSGVT	S		
18C16	MGYLTPKGRALKRMLDRIRRRREVKSGVT	S		
18C17	MKHSNSSRGRKTLWRALTLWLLMQSLKRT	S		

Figure 6.14. Sequences of the round 18 pool targeting L-P6.1. The starting pool is shown as a reference. Peptide sequences were grouped into three families. Peptides belonging to the 18C1 family contain a 12 base deletion resulting loss of four amino acids, indicated by (-). Mutations are highlighted in bold text.

Representative sequences from each family (18C1, 18C5, and 18C17) were synthesized as ^{35}S -labeled fusions, and tested for binding to immobilized L-P6.1 (Figure 6.15). The results show that all three families bind immobilized L-P6.1, however the sequences possess different binding affinities. 18C1 showed the highest affinity,

explaining why it came to dominate the round 18 pool. Surprisingly, 18C5 showed much less affinity to L-P6.1 than 18C17 and 18C1; the frequency it appeared in the round 18 pool implied that it should have higher affinity. These data suggest that each family binds L-P6.1 utilizing different contacts to the RNA. Additionally, there is an inverse correlation between the net charge (+6, +9, +7 for 18C1, 18C5, and 18C17, respectively) and affinity for LP6.1, implying that contacts other than ionic interactions are important.

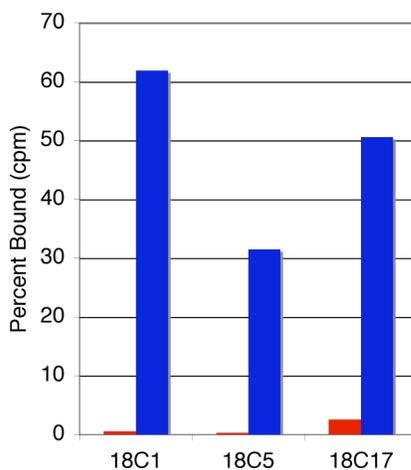


Figure 6.15. Binding of individual clones from round 18 of the L-P6.1 selection. Sequences representative of the three peptide families were synthesized as ^{35}S -labeled fusions and tested against immobilized L-P6.1. Fusions were also bound to matrix lacking target (red bars) to show binding was specific for L-P6.1.

We constructed a number of deletion constructs in order to define the minimal binding domain of each peptide. The peptides are predicted to adopt α -helical conformations (36) and these predicted helical regions were used to design deletion constructs. The deletion constructs were synthesized as ^{35}S -labeled fusions and several of the constructs showed little binding affinity toward immobilized L-P6.1 (Figure 6.16).

For peptide 18C1, nearly the entire randomized region is necessary for binding, while 18C17 shows reduced affinity if N- or C-terminal regions are deleted.

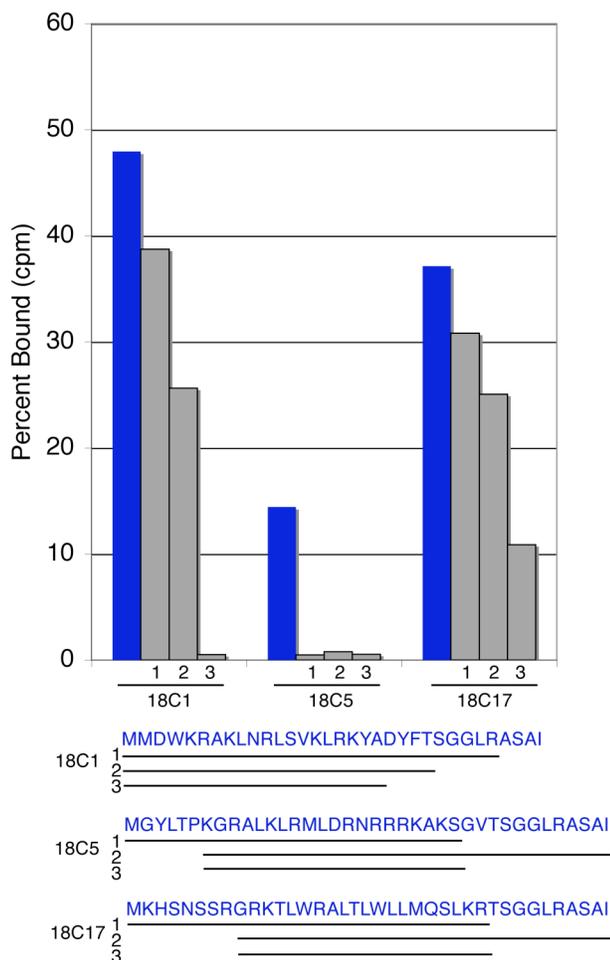


Figure 6.16. Deletion mapping of round 18C clones. Deletion constructs were constructed by PCR and tested for binding to immobilized L-P6.1. Black lines represent amino acids included in each respective construct. Methionine was added to constructs containing N-terminal deletions in order to initiate translation.

Conclusions

The peptides we have selected are highly basic, arginine rich, and are all predicted to form α -helicies. Our data support the notion that arginine is an important amino acid

for RNA recognition, both for its charge complementarity to RNA and its ability to form hydrogen bonds. While helical prediction programs are not always accurate, the fact that all sequences are predicted to be helical may stem from the favorable helical propensities of Arg and Lys. Other arginine rich sequences adopt α -helices (Rev (37)) or β -hairpins (Tat (38)), however sequences that recognize RNA hairpins are helical (39-43). It is therefore possible that α -helical structures are a general method to recognize RNA hairpins.

It is surprising that only a few solutions to a particular target were found. Only one in 10^{13} peptides possessed the ability to specifically bind the immobilized RNA targets with high affinity. Two selections using different random libraries resulted in the same consensus sequence, suggesting that very few peptide solutions exist for recognition of D-P6.1. The large number of selection cycles required for our experiments most likely reflects the relative paucity of solutions in the initial library. While it is probable that there are many more sequences in the selection pools, it is likely that few of these possess affinities comparable to the peptides we have selected since they were selected against. It is important to note that the peptides we have discovered would not have been found by phage display since a maximum of 10^9 sequences can be sieved by phage display.

It might seem, to a first approximation, that many RNA binding peptides should exist in the initial library since any sequence with basic residues should bind with high affinity (44-46). However, many of these sequences lack *specificity* and were removed by addition of tRNA competitor. Additionally, an inverse correlation is seen between binding affinity and total net charge of the selected peptides suggesting that other, non-electrostatic interactions are also important for RNA binding due to contributions to

overall affinity and specificity. Indeed, hydrophobic contacts that contribute to specificity are often found in other RNA-peptide or RNA-protein complexes (47,48).

In contrast with the selections described here, previous experiments targeting the *boxB* RNA resulted in selection for a diverse collection of RNA binding peptides (13). The difference in the number of selected solutions is probably due to the use of a constant region in the *boxB* selection that possessed micromolar affinity. The constant region probably increased the number of solutions in the starting pool, resulting in more sequences at the end of the selection. Indeed, the selection for *boxB* converged in fewer rounds than most of the selections described here. Additionally, mutagenesis and reselection of winning sequences described here against L-P6.1 and the L-HCV III_d targets resulted in many more independent solutions (Chapter 7).

Our results show that it is possible to excise a domain from a biologically important RNA and isolate high affinity peptides that bind this domain. While we have targeted RNA hairpins, natural peptides bind a variety of RNA structures (47, 48) and it is therefore likely that high affinity ligands to many other folded RNAs can be isolated from random peptide libraries. Since RNA constructs often fold into similar structures when removed from, or in the context of larger RNAs, our peptides probably bind to CR7 and P6.1 in the context of full-length hTR. Lastly, based on the *in vitro* binding assays, our peptides probably possess nanomolar affinities. Future experiments will address the ability of our peptides to inhibit telomerase activity.

Materials and Methods:

Construction of RRXRR Library

Single stranded DNA (5'-GGCGCAGCTGTTTCTGAGCTGGTAGCCGGCSNNS NNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNS21S21SNNS21S21SNNSNNSNNSNNSNNSNNS CATTGTAATTGTAAATAGTAATTGTCCC where N= an equimolar mixture of A,C,G,T; S= an equimolar mixture of G,C; 1=35:10:5:50 of T:C:A:G; 2=10:70:10:10 of T:C:A:G) was synthesized at the Keck Oligo Facility at Yale (<http://keck.med.yale.edu/oligos/>). The composition of each mixed base was adjusted to correct for the different coupling rates of each base (N= 3:3:2:2 for A:C:G:T; S= 3:2 for C:G; 1=23:10:5:33 for T:C:A:G; 2= 66:100:100:466 for T:C:A:G). The synthetic DNA was purified by 8% urea-PAGE and electroeluted. PCR conditions were optimized and were: 10 mM Tris-HCl, pH 9.0 at room temperature, 50 mM KCl, 2.5 mM MgCl₂, 0.1% Triton X-100 (Sigma), 4 pmol of template, 3 μM of the 42.108 5' primer (5'-TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA-3'), 3 μM of the 3' primer 3RRXRR (5'-GGCGCAGCTGTTTCTGAGCTGGTA-3'), and 0.2 mM each of dATP, dTTP, dCTP, and dGTP. PCR reactions were performed in 0.65 mL thick-walled tubes (final volume 200 μL, no heated lid on the thermocycler), heated to 95 °C before Taq was added, and cycled at 95 °C for two minutes followed by six cycles of two minutes at 95 °C, two minutes at 53 °C, and four minutes at 72 °C. A total of 24 mL of PCR was performed yielding 510-1200 μg of DNA product as determined by quantification against a mass ladder on an agarose gel. The PCR reactions were

combined, phenol extracted, ethanol precipitated, and resuspended in 1.2 mL of 1X TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0).

The 5' primer (42.108) was ³²P-labeled with polynucleotide kinase (PNK) and γ -³²P ATP. One cycle of PCR was performed under the conditions described above, except the 3' primer was omitted and template: 5' primer molar ratios were varied from 2:1, 5:1, 10:1, and 20:1. Roughly 20% of the input template was extendable (data not shown).

Eight hundred picomoles of PCR product were used in eight, one mL transcription reactions for more than 2 hours at 37 °C. Transcription reactions were 40 mM HEPES-KOH, pH 7.5, 2 mM spermidine, 40 mM DTT, 25 mM MgCl₂, 1X RNASecure (Ambion) and 4 mM each of UTP, GTP, ATP, and CTP (Sigma) at pH 8.0. The reactions were stopped by addition of 100 μ L of 0.5 M EDTA, pH 8.0, and were phenol extracted and ethanol precipitated. The RNA was purified by 8% urea-PAGE, electroeluted, and ethanol precipitated.

The RNA was ligated to pF30P (5' pdA21-[C9]3-dAdCdC-puromycin-3', where p is a 5' phosphate and C9 is spacer phosphoramidite 9 (Glen research)) using the splint RRXRR splint (5'- TTTTTTTTTTTNGGCGCAGCTGT-3', where N is 70% T and 10% of A, C, and G) and T4 DNA Ligase (NEB) in T4 DNA ligase buffer (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 1 mM ATP, 10 mM DTT, and 25 μ g/mL Bsa). The ligation reaction consisted of RNA, splint, and pF30P at a molar ratio of 1:1.5:1, respectively. After two hours at room temperature, the ligation reaction was purified by denaturing PAGE, electroeluted, and ethanol precipitated. The RNA and ligated template were quantified by absorbance at 260 nm and using the biopolymer calculator (<http://paris.chem.yale.edu/extinct.html>).

Rabbit reticulocyte lysate was prepared according to the method of Jackson and Hunt (49) and translation and fusion formation efficiencies were determined to be comparable to those obtained in a commercial rabbit reticulocyte lysate (Novagen). Fusion formation efficiencies were determined to be ~20% by translation of a ^{32}P -labeled ligated template and visualization on a tricine gel or urea-PAGE. A total of 20 mL of the homemade lysate and 5 mL of the commercial lysate (Novagen) were used to translate 10,000 pmol of ligated template (0.4 μM final concentration). Both lysates contained 20 mM HEPES-KOH, pH 7.6, 100 mM KCl, 0.5 mM MgCl_2 , 2 mM DTT, 8 mM creatine phosphate, 25 μM of each amino acid except for Leu and Cys, which were each 12.5 μM . A small amount of ^{32}P -labeled ligated template was added so that purification yields could be determined. Translation was conducted for one hour at 30 °C, after which KCl and MgCl_2 were added to final concentrations of 500 mM and 50 mM, respectively.

In order to simplify sample purification, the lysate was processed in 2.5 mL batches. For each 2.5 mL of translation, dT purification was performed with 100 mg of dT cellulose (NEB) prewashed with Isolation buffer (IB, 100 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.2% Triton X-100) and 25 mL of IB. After one hour at 4 °C, the entire sample was added to a plastic disposable column (Biorad) containing a frit blocked with glass wool. The beads were washed with 20 mL (~20 column volumes) of IB and eluted with five, 0.5 mL fractions of ddH₂O. The sample was precipitated in the presence of linear acrylamide (Ambion) and reverse transcribed in a total volume of 400 μL . The reactions were then incubated with 800 pmol of annealed target immobilized on neutravidin agarose (Pierce), and the buffer adjusted with 36 μL of 2.5 M KCl, 2.4 μL of 5% Nonidet P-40 (NP-40), 3 μL of 20 mg/mL yeast tRNA (Roche), and 680 μL of ddH₂O. The beads

were transferred to a Spin-X column, washed three times with 700 μL of 1X N binding buffer (NBB, 10 mM HEPES-KOH pH 7.5, 0.5 mM EDTA, pH 8.0, 100 mM KCl, 1 mM MgCl_2 , 1 mM DTT, 0.01% (v/v) NP-40), + 50 $\mu\text{g}/\text{mL}$ tRNA, and transferred to a new eppendorf tube with 1 mL of 1X NBB + tRNA, and the supernatant removed.

One-tenth of the beads was used to determine the number of cycles required for amplification of the bound DNA. The PCR reactions were performed under conditions described above except that the cycle times were shortened to 30 seconds at 94 $^\circ\text{C}$, 30 seconds at 55 $^\circ\text{C}$, and 45 seconds at 72 $^\circ\text{C}$.

The yields of dT and ethanol precipitation ranged from 57-74% of input mRNA template recovered. The RT efficiencies were determined to be >90%. This resulted in roughly 3×10^{14} molecules that entered the selective step for the D-P6.1 selection (~ 3.1 copies of each individual sequence) and $\sim 3 \times 10^{14}$ molecules that entered the selective step for the D-CR7 selection (~ 2.9 copies of each individual sequence).

RRXRR Selection Rounds

Subsequent rounds of selection were similar to the first round except reaction volumes were reduced. Each round consisted of one mL of PCR, which was phenol extracted and ethanol precipitated, and resuspended in 200 μL of 1X TE buffer. Half of the DNA stock was used in a one mL *in vitro* transcription for >2 hours at 37 $^\circ\text{C}$. The mRNA was desalted with a NAP-25 column and used in a 250-500 μL ligation. Ligated mRNA was translated in a 100 μL reaction after which salts were added and the reaction incubated at room temperature for one hour or overnight at -20 $^\circ\text{C}$. The fusions were purified by dT-cellulose with one mL of IB and 100 μL of 25% (v/v) dT cellulose pre-equilibrated in IB.

The DNA was phenol extracted, ethanol precipitated, and resuspended in 1 mL of 1X TE buffer (10 mM Tris-HCl, pH 7.5, 1 mM EDTA, pH 8.0) yielding a solution of 370 ng of DNA/ μ L. DNA concentration was determined by comparing a dilution series to a mass ladder (100 bp ladder, NEB).

Roughly 670 pmol of DNA (\sim 73 μ g of DNA) were *in vitro* transcribed in a 6 mL volume for 3 hours at 37 °C. RNA was gel purified by 6% denaturing PAGE and electroeluted. RNA was ligated to pF30P using the splint X27Splint (5'-TTTTTTTTTTTNAATAGCGGATG-3', where N is 70% T and 10% of A, C, and G) and T4 DNA Ligase (NEB). The ligation reaction consisted of RNA, splint, and pF30P at a molar ratio of 1:1.5:1, respectively. After two hours at room temperature, the ligation reaction was purified by denaturing PAGE, and electroeluted, yielding 3,600 pmol of ligated product.

Ligated mRNA (3,000 pmol) was translated in a total volume of 7.5 mL (0.4 μ M template) as described previously (13), salt was added to facilitate fusion formation, and the reaction incubated overnight at -20 °C. A small amount of ³²P-labeled ligated template was translated so that product yields could be determined.

The translated material was purified with 350 mg of dT-cellulose (NEB) that had been prewashed with Isolation buffer (IB, 100 mM Tris-HCl, pH 8.0, 1 M NaCl, 0.2% Triton X-100). The dT purification was performed at 4 °C for one hour, after which the beads were washed with 20 mL of IB on three disposable columns (Biorad) containing frits blocked with glass wool. Bound fusions on each column were eluted with eight, 0.5 mL washes of 65 °C ddH₂O. All eluates were pooled, then ethanol precipitated in the presence of linear acrylamide, as described previously (14). The pellets were

resuspended in 1X RT buffer (50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM MgCl₂, 10 mM DTT, 0.25 mM each dNTP, 1 μM of the 3' X27S primer) filtered with a 0.45 μm Spin-X column (Costar, Corning), and the top of the column washed with one 150 μL and two 50 μL fractions of ddH₂O. This step was necessary to remove dT cellulose that had been carried over from the dT purification step. All fractions were combined and the buffer adjusted to 1X RT in a final volume of 1,212 μL. The reactions were heated to 65 °C for 5 minutes, cooled on ice, and Superscript II (Invitrogen) was added. The samples were incubated at 42 °C for 1 hour then heated to 65 °C for 5 minutes to deactivate the Superscript II. The overall yield of material carried through dT purification and ethanol precipitation was 43-54%. Fusion efficiency was measured to be 15% of input template by translation of a ³²P-labeled ligated template (data not shown) while RT efficiency was measured to be >90%. Half the sample was used for selection against L-P6.1 while the other half was bound to D-P6.1 (C. Ueda, unpublished observations). Based on these yields, 5-6 x 10¹³ molecules entered the selective step for each RNA target, of which 1-2 x 10¹³ were independent sequences.

L-P6.1 (600 pmol) was annealed in 1X selection buffer (SB, 10 mM Tris-HCl, pH 7.5, 150 mM KCl, 0.5 mM EDTA, pH 8.0, 0.01% (v/v) Tween-20, 1 mM DTT) by heating the sample to 90 °C for one minute, then cooling rapidly on ice. Neutravidin agarose (112.5 μL of beads) was pre-equilibrated in 1X SB. Annealed L-P6.1 was added to the beads and the mixture incubated at room temperature for 30 minutes after which the supernatant was removed.

The buffer of the reverse transcription reaction was adjusted to selection conditions by the addition of 54 μL of 2.5 M KCl, 1.8 μL of 10% Tween-20, 9 μL of 10

mg/mL tRNA, 2 μ L of EDTA, pH 8.0, and 1,000 μ L of H₂O. The entire reaction was added to the immobilized L-P6.1 on neutravidin beads and incubated at 4 °C for 1.5 hours. Twenty microliters of 2 mM biotin in 1X PBS (4.3 mM Na₂HPO₄-7H₂O, 1.4 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl) was added and the reaction incubated at 4 °C for an additional 30 minutes. The beads were transferred to a Spin-X column, washed three times with 700 μ L of 1X SB + 50 μ g/mL tRNA, and transferred to a new eppendorf tube with 1 mL of 1X SB + tRNA, and the supernatant removed.

One-tenth of the beads was used to determine the number of cycles required for amplification of the bound DNA. The PCR reactions were performed under conditions described above except that the cycle times were shortened to 30 seconds at 94 °C, 30 seconds at 55 °C, and 45 seconds at 72 °C. Six, nine, and eleven cycles of PCR were performed and the products run on a sodium borate-agarose gel (50). The remaining bound DNA was then amplified for 11 cycles in a total volume of 1 mL.

L-P6.1 Selection Rounds

The selection rounds for L-P6.1 using the X27S library are similar to those described for the RRXRR library, except as detailed below. One microliter of 100 μ M L-P6.1 was added to 50 μ L of binding buffer -tRNA, and annealed by heating to 90 °C and rapidly cooling on ice. The annealed RNA was added to 400 μ L of binding buffer and 50 μ L of 50:50 (v/v) neutravidin agarose slurry prewashed with binding buffer, and bound for one hour at room temperature. The supernatant was removed, and the reverse transcribed fusions, one mL of binding buffer with tRNA, and 10 μ L of 2 mM D-biotin in 1X PBS was added. The fusions were incubated at 4 °C for one hour after which the

slurry was transferred to a Spin-X column and washed five times with 700 μL fractions of binding buffer with tRNA. The beads were resuspended with 700 μL of binding buffer, transferred to a 1.7 mL eppendorf tube, and the supernatant removed. 75 μL of ddH₂O was added and 10 μL of the slurry used to determine the number of cycles required for PCR.

For rounds where the matrix changed, Ultralink streptavidin (Pierce) was used instead of neutravidin agarose. Rounds where the magnetic streptavidin matrix (Genovision) was used were similar to those described above, except that after the fusions were bound, the magnetic beads were washed four times with 1 mL of binding buffer with tRNA using a robotic liquid handling system (Kingfisher, ThermoLab Systems). The last wash was conducted using a hand-held magnet and the beads were resuspended in 100 μL of ddH₂O before PCR. For rounds where tRNA concentration was increased, high tRNA concentrations (500 $\mu\text{g}/\text{mL}$ or 5,000 $\mu\text{g}/\text{mL}$) were used in both the binding reaction and washes, except for the last few washes where tRNA concentration was decreased by ten-fold until it reached 50 $\mu\text{g}/\text{mL}$.

The matrix used for each round of selection is denoted in Figure 6.S.2.

***In Vitro* Binding Assay**

dT-purified ³⁵S-labeled fusions (~100,000 cpm) were incubated with 0.1-0.2 μM immobilized, annealed RNA target (50 μL of a 50/50 (v/v) slurry) in binding buffer with 50 $\mu\text{g}/\text{mL}$ tRNA for one hour at 4 °C. The samples were transferred to Spin-X columns

(Costar) and washed three times with 700 μ L of binding buffer. The supernatant, washes, and beads counted by scintillation.

Synthesis of RNA Constructs

For NMR and native gel shift experiments, RNA constructs were synthesized by *in vitro* transcription of partially ssDNA templates using T7 RNA Polymerase (34). The D-CR7 template (5'-GTCCCACAGCTCAGGGACTATAGTGAGTCGTATTACGAATT-3') and D-P6.1 template (5'-GGAGAGCCCAACTCTCCTATAGTGAGTCGTATTACGAATT-3') were used with T7 promoter strand (5'-AATTCGTAATACGACTCACTATA-3') for the transcription of D-CR7 and D-P6.1 RNAs, respectively. The RNA was gel purified by 20% denaturing PAGE, electroeluted, and ethanol precipitated. The electroeluted RNA was used directly for native gel experiments.

RNA used for NMR was dissolved in 2 M NaCl and dialyzed extensively against ddH₂O to remove any Tris present in the sample. RNA was then freeze-dried, and resuspended in NMR buffer (10 mM phosphate, pH 6.0, 50 mM NaCl) in H₂O/D₂O (90:10 (v/v)).

D-P6.1 (5'-photocleavable biotin (Glen research)-AAAGGAGAGUUGGGCUCUCC-3') and D-CR7 (5'-photocleavable biotin-AAAGUCCCUGAGCUGUGGGAC-3') were synthesized at the Caltech Oligo Facility. The 2'-silyl protecting groups were removed by addition of 360 μ L of triethylamine trihydrofluoride (Aldrich), incubation overnight at room temperature, and butanol precipitation. L-P6.1 (5'-GGAGAGUUGGGCUCUCCAAA-Biotin-3') was synthesized and supplied deprotected by Chemgenes. RNAs were purified by 20% denaturing PAGE and electroeluted.

Peptide Synthesis

The 8H1 peptide (NH₂-MNDARRNRKYLRVKRLRIQKMY-COOH) was synthesized using fluorenylmethyloxycarbonyl (Fmoc) chemistry on a 432A Synergy Peptide Synthesizer (Applied Biosystems). The peptide was deprotected in a trifluoroacetic acid (TFA)/1,2-ethanediol/thioanisole mixture (90:5:5, (v/v)) for 2 hours then washed with methylterbutyl ether and resuspended in ddH₂O. The crude peptide mixture was purified by reverse-phase HPLC using a C18 semiprep column (250x10 mm, Vydac). Buffer A was a 95:5:0.1 (v/v) mixture of water/acetonitrile/TFA while buffer B was a 10:90:0.1 (v/v) mixture of water/acetonitrile/TFA. A gradient from 0%-40% B was run in 120 minutes and fractions containing the peptide were pooled. Final purity (>95%) was assessed by analytical HPLC on an analytical C18 column (250x4.6 mm, Vydac) using buffer A and buffer B (gradient of 0-100% buffer B in 50 minutes). The peptide mass was confirmed by MALDI-TOF mass spectrometry. The peptide concentration was determined by absorbance of tyrosine using the biopolymer calculator (<http://paris.chem.yale.edu/extinct.html>).

Native Gel Electrophoresis

RNAs were annealed in running buffer (0.5X TBE (44.5 mM Tris base, 44.5 mM boric acid, pH 8.3, 1 mM EDTA) containing 10% glycerol) by heating to 90 °C and rapidly cooling on ice. RNAs were run on a 10% acrylamide gel in a thermostated gel box (Hoeffer) in 0.5X TBE at 4 °C. Roughly 5,000 picomoles of RNA were run on each lane in order for the RNA to be visualized by UV-shadowing.

NMR Spectroscopy

NMR spectra were collected at 25 °C as previously described (13, 15). RNAs were annealed before the NMR experiments by heating the sample to 90 °C and cooling rapidly on ice. The concentration of P6.1 was 1.5 mM while CR7 was ~0.5 mM.

Sequencing

Pool templates were PCR amplified, gel purified using the QiaQuick gel extraction kit (Qiagen), and cloned using a TOPO-TA cloning kit (Invitrogen). Single colonies were grown in LB broth containing 50-100 µg/mL Ampicillin, and purified using the QiaPrep Spin Miniprep kit (Qiagen). Sequences were analyzed by Sequencher.

Biacore Analysis

Biacore analysis was performed on a Biacore 2000 instrument at 25 °C. A streptavidin chip (Sensor Chip SA, Biacore) was preconditioned by three 50 µL injections of 40 mM NaOH, 2 M NaCl. Annealed P6.1 (5'-GGAGAGUUGGGCU CUCC-biotin) in HBS-EP buffer (10 mM HEPES-KOH pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% (v/v) Tween-20) was injected at a 1 µM concentration yielding ~1,000 RU of immobilized P6.1. 8H1 was diluted into running buffer (HBS-EP, 1 mM DTT, 50 µg/mL tRNA, 0.5 mg/mL Bsa) and injected at 99 nM, 33 nM, 11 nM, 3.67 nM, and 1.22 nM. The sensor chip was regenerated with injections of 2 M NaCl. Data were processed and analyzed using the BIAEvaluation software (Biacore).

References

1. Cech, T.R. Life at the end of the chromosome: telomeres and telomerase. (2000) *Angew Chem Int Ed Engl* **39**, 34-43.
2. Cong, Y.S., Wright, W.E. and Shay, J.W. Human telomerase and its regulation. (2002) *Microbiol Mol Biol Rev* **66**, 407-425, table of contents.
3. Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L., Coviello, G.M., Wright, W.E., Weinrich, S.L. and Shay, J.W. Specific association of human telomerase activity with immortal cells and cancer. (1994) *Science* **266**, 2011-2015.
4. Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W. and Weinberg, R.A. Creation of human tumour cells with defined genetic elements. (1999) *Nature* **400**, 464-468.
5. Counter, C.M., Hirte, H.W., Bacchetti, S. and Harley, C.B. Telomerase activity in human ovarian carcinoma. (1994) *Proc Natl Acad Sci U S A* **91**, 2900-2904.
6. Harrison, R.J., Gowan, S.M., Kelland, L.R. and Neidle, S. Human telomerase inhibition by substituted acridine derivatives. (1999) *Bioorg Med Chem Lett* **9**, 2463-2468.
7. Sun, D., Thompson, B., Cathers, B.E., Salazar, M., Kerwin, S.M., Trent, J.O., Jenkins, T.C., Neidle, S. and Hurley, L.H. Inhibition of human telomerase by a G-quadruplex-interactive compound. (1997) *J Med Chem* **40**, 2113-2116.

8. Strahl, C. and Blackburn, E.H. Effects of reverse transcriptase inhibitors on telomere length and telomerase activity in two immortalized human cell lines. (1996) *Mol Cell Biol* **16**, 53-65.
9. Asai, A., Oshima, Y., Yamamoto, Y., Uochi, T.A., Kusaka, H., Akinaga, S., Yamashita, Y., Pongracz, K., Pruzan, R., Wunder, E., Piatyszek, M., Li, S., Chin, A.C., Harley, C.B. and Gryaznov, S. A novel telomerase template antagonist (GRN163) as a potential anticancer agent. (2003) *Cancer Res* **63**, 3931-3939.
10. Norton, J.C., Piatyszek, M.A., Wright, W.E., Shay, J.W. and Corey, D.R. Inhibition of human telomerase activity by peptide nucleic acids. (1996) *Nat Biotechnol* **14**, 615-619.
11. Pitts, A.E. and Corey, D.R. Inhibition of human telomerase by 2'-O-methyl-RNA. (1998) *Proc Natl Acad Sci U S A* **95**, 11549-11554.
12. Dominick, P.K., Keppler, B.R., Legassie, J.D., Moon, I.K. and Jarstfer, M.B. Nucleic acid-binding ligands identify new mechanisms to inhibit telomerase. (2004) *Bioorg Med Chem Lett* **14**, 3467-3471.
13. Barrick, J.E., Takahashi, T.T., Ren, J., Xia, T. and Roberts, R.W. Large libraries reveal diverse solutions to an RNA recognition problem. (2001) *Proc Natl Acad Sci U S A* **98**, 12374-12378.
14. Barrick, J.E., Takahashi, T.T., Balakin, A. and Roberts, R.W. Selection of RNA-binding peptides using mRNA-peptide fusions. (2001) *Methods* **23**, 287-293.
15. Xia, T., Frankel, A., Takahashi, T.T., Ren, J. and Roberts, R.W. Context and conformation dictate function of a transcription antitermination switch. (2003) *Nat Struct Biol* **10**, 812-819.

16. Takahashi, T.T., Austin, R.J. and Roberts, R.W. mRNA display: ligand discovery, interaction analysis and beyond. (2003) *Trends Biochem Sci* **28**, 159-165.
17. Burd, C.G. and Dreyfuss, G. Conserved structures and diversity of functions of RNA-binding proteins. (1994) *Science* **265**, 615-621.
18. Futaki, S., Suzuki, T., Ohashi, W., Yagami, T., Tanaka, S., Ueda, K. and Sugiura, Y. Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. (2001) *J Biol Chem* **276**, 5836-5840.
19. Wender, P.A., Mitchell, D.J., Pattabiraman, K., Pelkey, E.T., Steinman, L. and Rothbard, J.B. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: peptoid molecular transporters. (2000) *Proc Natl Acad Sci U S A* **97**, 13003-13008.
20. Williams, K.P., Liu, X.H., Schumacher, T.N., Lin, H.Y., Ausiello, D.A., Kim, P.S. and Bartel, D.P. Bioactive and nuclease-resistant L-DNA ligand of vasopressin. (1997) *Proc Natl Acad Sci U S A* **94**, 11285-11290.
21. Schumacher, T.N., Mayr, L.M., Minor, D.L., Jr., Milhollen, M.A., Burgess, M.W. and Kim, P.S. Identification of D-peptide ligands through mirror-image phage display. (1996) *Science* **271**, 1854-1857.
22. Eckert, D.M., Malashkevich, V.N., Hong, L.H., Carr, P.A. and Kim, P.S. Inhibiting HIV-1 entry: discovery of D-peptide inhibitors that target the gp41 coiled-coil pocket. (1999) *Cell* **99**, 103-115.

23. Antal, M., Boros, E., Solymosy, F. and Kiss, T. Analysis of the structure of human telomerase RNA *in vivo*. (2002) *Nucleic Acids Res* **30**, 912-920.
24. Chen, J.L., Blasco, M.A. and Greider, C.W. Secondary structure of vertebrate telomerase RNA. (2000) *Cell* **100**, 503-514.
25. Chen, J.L., Opperman, K.K. and Greider, C.W. A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. (2002) *Nucleic Acids Res* **30**, 592-597.
26. White, L.K., Wright, W.E. and Shay, J.W. Telomerase inhibitors. (2001) *Trends Biotechnol* **19**, 114-120.
27. Lukowiak, A.A., Narayanan, A., Li, Z.H., Terns, R.M. and Terns, M.P. The snoRNA domain of vertebrate telomerase RNA functions to localize the RNA within the nucleus. (2001) *Rna* **7**, 1833-1844.
28. Allain, F.H. and Varani, G. Structure of the P1 helix from group I self-splicing introns. (1995) *J Mol Biol* **250**, 333-353.
29. Leeper, T., Leulliot, N. and Varani, G. The solution structure of an essential stem-loop of human telomerase RNA. (2003) *Nucleic Acids Res* **31**, 2614-2621.
30. Ueda, C.T. and Roberts, R.W. Analysis of a long-range interaction between conserved domains of human telomerase RNA. (2004) *Rna* **10**, 139-147.
31. Mitchell, J.R. and Collins, K. Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. (2000) *Mol Cell* **6**, 361-371.
32. Rist, M. and Marino, J. Association of an RNA kissing complex analyzed using 2-aminopurine fluorescence. (2001) *Nucleic Acids Res* **29**, 2401-2408.

33. Beattie, T.L., Zhou, W., Robinson, M.O. and Harrington, L. Functional multimerization of the human telomerase reverse transcriptase. (2001) *Mol Cell Biol* **21**, 6151-6160.
34. Milligan, J.F. and Uhlenbeck, O.C. Synthesis of small RNAs using T7 RNA polymerase. (1989) *Methods Enzymol* **180**, 51-62.
35. Tan, R. and Frankel, A.D. Structural variety of arginine-rich RNA-binding peptides. (1995) *Proc Natl Acad Sci U S A* **92**, 5282-5286.
36. Rost, B. and Sander, C. Prediction of protein secondary structure at better than 70% accuracy. (1993) *J Mol Biol* **232**, 584-599.
37. Battiste, J.L., Mao, H., Rao, N.S., Tan, R., Muhandiram, D.R., Kay, L.E., Frankel, A.D. and Williamson, J.R. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. (1996) *Science* **273**, 1547-1551.
38. Puglisi, J.D., Chen, L., Blanchard, S. and Frankel, A.D. Solution structure of a bovine immunodeficiency virus Tat-TAR peptide-RNA complex. (1995) *Science* **270**, 1200-1203.
39. Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S. and Rosch, P. Antitermination in bacteriophage lambda. The structure of the N36 peptide-*boxB* RNA complex. (2000) *Eur J Biochem* **267**, 2397-2408.
40. Legault, P., Li, J., Mogridge, J., Kay, L.E. and Greenblatt, J. NMR structure of the bacteriophage lambda N peptide/*boxB* RNA complex: recognition of a GNRA fold by an arginine-rich motif. (1998) *Cell* **93**, 289-299.
41. Cilley, C.D. and Williamson, J.R. Structural mimicry in the phage phi21 N peptide-*boxB* RNA complex. (2003) *Rna* **9**, 663-676.

42. Cai, Z., Gorin, A., Frederick, R., Ye, X., Hu, W., Majumdar, A., Kettani, A. and Patel, D.J. Solution structure of P22 transcriptional antitermination N peptide-*boxB* RNA complex. (1998) *Nat Struct Biol* **5**, 203-212.
43. Faber, C., Scharpf, M., Becker, T., Sticht, H. and Rosch, P. The structure of the coliphage HK022 Nun protein-lambda-phage *boxB* RNA complex. Implications for the mechanism of transcription termination. (2001) *J Biol Chem* **276**, 32064-32070.
44. Mascotti, D.P. and Lohman, T.M. Thermodynamics of oligoarginines binding to RNA and DNA. (1997) *Biochemistry* **36**, 7272-7279.
45. Mascotti, D.P. and Lohman, T.M. Thermodynamics of single-stranded RNA binding to oligolysines containing tryptophan. (1992) *Biochemistry* **31**, 8932-8946.
46. Mascotti, D.P. and Lohman, T.M. Thermodynamics of single-stranded RNA and DNA interactions with oligolysines containing tryptophan. Effects of base composition. (1993) *Biochemistry* **32**, 10568-10579.
47. Draper, D.E. Protein-RNA recognition. (1995) *Annu Rev Biochem* **64**, 593-620.
48. Draper, D.E. Themes in RNA-protein recognition. (1999) *J Mol Biol* **293**, 255-270.
49. Jackson, R.J. and Hunt, T. Preparation and use of nuclease-treated rabbit reticulocyte lysates for the translation of eukaryotic messenger RNA. (1983) *Methods Enzymol* **96**, 50-74.
50. Brody, J.R. and Kern, S.E. Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis. (2004) *Biotechniques* **36**, 214-216.